**Prognostic value of first recorded presentation of breathlessness for future COPD, asthma, IHD, and mortality: a cohort study using UK national primary care database**

**Running headline: Dyspnoea in primary care**

Ying Chen1 (PhD), Richard Hayward1 (PhD), Carolyn Chew-Graham1,3 (MD), Richard Hubbard2 (MD), Peter Croft1 (MD), Keith Sims4, Kelvin P Jordan1 (PhD)

1 Research Institute for Primary Care and Health Sciences, David Weatherall Building, Keele University, Staffordshire, ST5 5BG; 2 Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Nottingham, UK; 3 NIHR Collaboration for Leadership in Applied Health Research and Care West Midlands, Keele, UK; 4 North Staffordshire branch, ‘Breathe Easy’ British Lung Foundation Support Group, UK.

Ying Chen, corresponding author (Research Institute for Primary Care & Health Sciences, Keele University, Staffordshire UK, ST5 5BG; y.chen1@keele.ac.uk; Tel:  +44(0)1782 733990).

**Abstract**

**Background:** Breathlessness is a common presentation in primary care, but GPs may not always record a diagnosis.

**Aim:** To assess long-term risk of diagnosed COPD, asthma and IHD, and mortality in patients with undiagnosed breathlessness.

**Design and Setting:** Matched cohort study set within the Clinical Practice Research Datalink.

**Method**: Adults with first recorded breathlessness between 1997–2010 and no prior diagnostic or prescription record for IHD or a respiratory disease (‘exposed’ cohort) were matched to an ‘unexposed’ cohort. Analyses were adjusted for socio-demographic and co-morbid characteristics.

**Results:** 75,698 ‘exposed’ patients were followed for a median of 6.1 years, and more than a third subsequently received a diagnosis of COPD, asthma or IHD. In those who remained undiagnosed after six months, there were increased long-term risks of all three diagnoses compared to the ‘unexposed’ cohort. Adjusted hazard ratios for COPD ranged from 8.6 (95%CI 6.8, 11.0) between 6-12 months to 2.8 (2.6, 3.0) after 36 months (asthma 11.7 to 4.3; IHD 3.0 to 1.6). Risk of long-term diagnosis remained higher in the ‘exposed’ cohort with no relevant prescription in the first 6 months, and approximately half of all future diagnoses were made in such patients. Risk of earlier mortality (all-cause and disease-specific) was higher in the ‘exposed’ cohort.

**Conclusion:** Undiagnosed breathlessness can be an indicator of later COPD, asthma and IHD, and is associated with earlier mortality. There is potential for improved identification at first presentation in primary care of those at high risk of future disease, with careful assessment, appropriate intervention and pro-active follow-up and monitoring.

**Keywords:** Dyspnoea, Chronic Obstructive Pulmonary Disease, Asthma, Ischemic Heart Disease, Primary Health Care, Electronic Health Records.

**How this fit in**

The prospective picture of what happens in the future to patients who present with initial breathlessness in primary care but not given a diagnostic label for their symptom was not known. This study shows that future disease burden of COPD, asthma and IHD is high and earlier mortality is increased in patients presenting undiagnosed breathlessness. In those who receive no relevant prescription in the first 6 months, the increased risk of future diagnoses of COPD, asthma and IHD remains, which may represent a missed opportunity for early management. Our findings suggest the potential for more and detailed investigation and monitoring to better confirm or rule out COPD, asthma and IHD, and for early intervention among patients presenting with breathlessness.

**Introduction**

Early diagnosis of chronic disease, with the possibility of early intervention, could deliver significant health benefits.1-5 For example, early recognition of chronic obstructive pulmonary disease (COPD) and intervention with smoking cessation could have major health benefits for the patient,6-8 whilst timely intervention with appropriate medication in asthma alleviates symptoms and reduces emergency admissions in adolescents and adults.9

In UK primary care, problems, including disease labels, are generally recorded using the ‘Read’ system of codes.10 The primary care clinician may record a symptom in medical records when patients presenting with this symptom at first presentation, if he/she is unsure of the underlying diagnosis, or believe the symptom represents a minor or self-limiting illness. The act of labelling a consultation with a symptom rather than with the underlying (unknown) diagnosis may affect the management and outcome of that patient. The use of a diagnostic record may flag up specific management strategies. For example, coding a patient as heart failure or ischaemic heart disease (IHD) will mean that the clinician is likely to initiate specific treatments and review patients in line with single disease guidelines, and, in England, the Quality and Outcomes Framework (QOF).11

Breathlessness is a common respiratory symptom. The prospective picture of what happens in the future to patients who present with initial breathlessness in primary care but not given a diagnostic label for their symptom was not known. The objectives of this study were to assess long-term risk of diagnosed COPD, asthma and IHD, and all-cause and disease-specific mortality in patients with a new recorded symptom in primary care of breathlessness, compared to patients without such a symptom.

**Method**

*Database*

This was a cohort study set within the Clinical Practice Research Datalink (CPRD).12 Diagnoses recorded in CPRD have been validated in a wide range of diseases, including respiratory and circulatory conditions.13,14

For this study, practices included (*n*=360, all from England) were those with linkage to hospital inpatient information (Hospital Episode Statistics, HES), neighbourhood deprivation (Index of Multiple Deprivation 2010, IMD2010) and mortality (via the Office of National Statistics, ONS) data. Similar characteristics in respect of patient demographics, years of follow-up and prescribing behaviour have been shown in practices with and without such linkages.15

*Participant identification*

All individuals aged 18 years and over with a first coded record of a breathlessness symptom during 1997–2010 were identified as the ‘exposed’ cohort. Prior to the first breathlessness code, patients included in this ‘exposed’ cohort had no diagnostic record of COPD, asthma, IHD or other respiratory disease. Furthermore, patients had no prescription record of medications for COPD, asthma or IHD in the 2 years prior to the first breathlessness symptom. The date of first coded breathlessness symptom was defined as the index date.

An ‘unexposed’ cohort was matched 1:1 by year of birth, gender and practice to the ‘exposed’ cohort. The ‘unexposed’ cohort did not have a prior recorded breathlessness symptom, a prior diagnosis of COPD, asthma, IHD or other respiratory disease, or prescription record of medications for COPD, asthma, or IHD for the 2 years prior to the index date of their matched ‘exposed’ patient. They were given the same index date as their matched exposed patient.

‘’The matched ‘exposed’ and ‘unexposed’ cohorts were followed forward from index date until earliest of 31st March 2014 or the date they no longer contributed to CPRD (e.g. due to death, leaving the practice, the practice leaving CPRD).

*Outcomes*

New diagnoses of COPD, asthma and IHD were defined as the first recording of these conditions in either CPRD or the linked hospital inpatient data within the follow-up period. They were analysed separately so a patient could have multiple outcomes. COPD, asthma and IHD were identified using Read codes and these Read codes were mapped to International Classification of Diseases v10 codes in the linked HES dataset to enable outcome identification from inpatient hospital information as well as primary care. Read codes used are available from [www.keele.ac.uk/mrr](http://www.keele.ac.uk/mrr).

Mortality information, including date and cause, was obtained from the linked ONS data. COPD and IHD-related as well as all-cause mortality were determined. Asthma-related mortality was not studied because of its rarity.

*Covariates*

Covariates thought to be potentially related to both the recording of exposure symptoms and outcomes were year of index date, age, gender, geographical region, smoking status, alcohol consumption, body mass index (BMI), neighbourhood deprivation, other recorded respiratory symptoms (cough, sputum), recorded inflammatory conditions (rheumatoid arthritis, gout, polymyalgia rheumatica, inflammatory bowel disease, systemic lupus erythematosus, spondyloarthritis), common mental health disorders (depression, anxiety), diabetes, non-specific chest pain, musculoskeletal pain, and the total number of co-morbid conditions.

Smoking and alcohol information were classified as ever smoked/drunk alcohol, never or missing, based upon data recorded prior to index date. The missing category was included to maximise the numbers in the analysis. The BMI value used was the most recent record before index date, and was grouped into <25, ≥25 kg/m2 (overweight), or missing. Neighbourhood deprivation level was based on IMD2010, a weighted measure of deprivation across seven domains for geographical areas.16 This was categorised based on quintile score (1 least, 5 most deprived). Repeating analyses excluding patients with missing data on covariates gave similar results.

Cough, sputum and the specific co-morbidities were defined through Read code extracted from patients’ consultation records in the 2-year period prior to index date. The total number of prescriptions for drugs recorded under different British National Formulary sections in the 2-year period prior to index date was used as a surrogate measure of the number of co-morbid conditions at baseline. This has been shown to be as predictive of health outcomes as more complicated comorbidity measures.17

*Primary care management*

As a measure of primary care management, prescribed medications for COPD, asthma and IHD in the first six months after the index date (and before any relevant diagnosis) were identified. A six-month period was thought to be a sufficient time window to ensure that a diagnosis resulting from the initial symptom presentation found its way into the record, for example, in order to allow time for relevant investigations.

Medications relevant to COPD, asthma and IHD were identified through consensus of two academic primary care physicians (RH and CC-G). Two lists were compiled: 1) drugs used only for management of COPD/asthma or IHD (the narrow list), and 2) drugs used for COPD/asthma or IHD and possibly used in other conditions (the broad list). It was not possible to determine separate medication lists for COPD and asthma as these conditions often coexist in the same patient.18,19 The narrow list was used to exclude individuals with previous evidence of COPD, asthma or IHD before index date when identifying the study cohorts (as described above). The broad list was used to determine whether patients received medications relevant to COPD/asthma or IHD after the index date within the ‘exposed’ cohort. The broad list therefore generates a conservative estimate of patients who did not receive any relevant management. Medication lists are available from [www.keele.ac.uk/mrr](http://www.keele.ac.uk/mrr).

*Statistical analysis*

The rates of new diagnosis of COPD, asthma, IHD, all-cause and disease-specific mortality per 10,000 person-years at risk were determined in both ‘exposed’ and ‘unexposed’ cohorts, stratified by time from index date.

We defined the first six months after index date as the ‘diagnostic window’. We expected the risk of future diagnosis to vary by length of time from index date, and so we split the follow-up time from six months onwards into three periods: 6-12 months, 12-36 months, and more than 36 months after index date in those with no diagnosis by the start of the relevant period. Cox proportional hazards regression was used to produce both unadjusted and adjusted hazard ratios (HR) for each of the three disease outcomes separately, estimating the excess risk of each outcome diagnosis associated with a first presentation of breathlessness symptom. The validity of the proportional-hazards assumption was tested using Schoenfeld residuals and deemed adequate for the exposure variable. In the final models, all covariates were included as confounding factors for adjustment. We also stratified by smoking status (since smoking is a major risk factor for the symptom and disease outcomes) as a sensitivity analysis.

Within the ‘exposed’ cohort, patients who received a relevant prescription in the first six months after index date (and before the date of relevant diagnosis if any) were identified. All models were repeated to assess the risk of long-term diagnosis in those without any initial relevant prescription or diagnosis in the first six months, compared to the ‘unexposed’ cohort.

The risks of earlier all-cause mortality between the ‘exposed’ and ‘unexposed’ cohorts were compared using Cox regression models, with diagnosis (the earliest diagnosis of COPD, asthma or IHD) treated as a time-varying covariate. The risks of future disease-specific (COPD or IHD) mortality were assessed using competing-risk models with death from other causes as competing events.

All analyses were performed using STATA 14.2.

**Results**

*Baseline characteristics*

There were 75,698 patients included in each of the ‘exposed’ (recorded first breathlessness symptom) and ‘unexposed’ cohorts (43% male; median age 60 (IQR 46, 72)). The median follow-up times were 6.1 (IQR 3.5, 9.4) and 6.5 (IQR 3.9, 9.7) years in the ‘exposed’ and ‘unexposed’ cohorts, respectively. Characteristics of the two cohorts are given in Table 1 and (for individual comorbidities) Supplementary Table 1.

*Risk of COPD, asthma, and IHD diagnosis*

During the whole follow-up, 35% of patients in the ‘exposed’ cohort were diagnosed with at least one of COPD, asthma or IHD, whereas it was 11% in the ‘unexposed’ cohort (Supplementary Table 2). Kaplan-Meier curves show the higher rates of COPD, asthma and IHD diagnoses during the follow-up in the ‘exposed’ cohort compared to those in the ‘unexposed’ cohort (Figure 1).

4% of the ‘exposed’ cohort received a diagnosis of COPD (6% asthma, 6% IHD) in the initial six months (diagnostic window). Between 0.1% and 0.4% of the ‘unexposed’ cohort received one of these diagnoses in this 6-month period (Table 2).

In the long-term (>6 months), incidence of COPD in the ‘exposed’ cohort was 139 per 10,000 person-years (asthma 168/10,000; IHD 203/10,000), and in the ‘unexposed’ cohort it was 37/10,000 person-years (asthma 27/10,000; IHD 107/10,000) (Table 2). Long-term excess risk of all 3 diagnoses for those in the ‘exposed’ cohort was evident. This risk reduced over time but remained statistically significant. For COPD the adjusted HR ranged from 8.6 (95% CI 6.8, 11.0) between 6 and 12 months after index date to 2.8 (95% CI 2.6, 3.0) after 36 months. Adjusted HR for asthma ranged from 11.7 (95% CI 9.4, 14.6) to 4.3 (95% CI 3.9, 4.6), and for IHD from 3.0 (95% CI 2.7, 3.4) to 1.6 (95% CI 1.5, 1.7) (Table 2).

The findings were similar in both strata when stratifying by smoking status (Supplementary Table 3).

*Risk of diagnosis in those with no management within initial six months*

Within the first 6 months after index date (and before any relevant diagnosis), 35% of the ‘exposed’ cohort received COPD/asthma medication, and 34% received IHD medication. In those who received a COPD diagnosis at any time during follow-up, 53% received a relevant prescription in the first six months and prior to a diagnosis, compared with 32% in those without a COPD diagnosis (71% vs. 28% for asthma; 56% vs. 30% for IHD) (Table 3).

Patients with no relevant prescribing in the first six months after index date still had an increased long-term risk for all 3 diagnoses compared to the ‘unexposed’ cohort, although at lower risk than those who had been prescribed relevant medication (Table 4). This group of patients consisted of 54% of all those subsequently diagnosed with COPD during the long-term (>6 months) follow-up (40% of those diagnosed with asthma; 43% of those diagnosed with IHD).

*Risk of all-cause and disease-specific mortality*

During the whole follow-up, higher rates of all-cause (289 per 10,000 person-years), COPD (22/10,000) and IHD-related mortality (69/10,000) were observed in the ‘exposed’ cohort compared to the ‘unexposed’ cohort (all-cause 155/10,000; COPD-related 4/10,000; IHD-related 32/10,000).

In the long-term (>6 months), patients with an initial recorded breathlessness symptom had increased excess risk of all-cause (adjusted HR 1.2 95% CI (1.1, 1.3)), COPD (adjusted HR 1.8 95% CI (1.5, 2.1)) and IHD-related (adjusted HR 1.2 95% CI (1.1, 1.3)) mortality even after adjustment for later diagnosis (Table 5). Mortality risks stratified by age groups (18-59 and 60+) are shown in Supplementary Table 4.

**Discussion**

This study of over 150,000 primary care patients showed an increased long-term risk of COPD, asthma, and IHD diagnosis, and of earlier mortality among patients presenting with initially undiagnosed symptom of breathlessness. Over a median of 6 years, more than a third of these patients subsequently received a diagnosis of COPD, asthma or IHD.

The high risk of diagnoses in the first six months is likely to reflect the diagnostic period for the initial presentation of breathlessness to allow time for tests and monitoring of symptoms. However, the majority of patients who received a diagnosis of COPD, asthma or IHD did so after this six month period, which may imply a delay in diagnosis. The higher long-term risks of subsequent diagnosis of COPD, asthma or IHD and of mortality suggest the important prognostic value of this symptom, with one in ten patients receiving a diagnosis more than three years after their initial symptom presentation.

In the first six months after presenting with an undiagnosed breathlessness symptom, one third of the patients received medications related to COPD/asthma and one third IHD medications, which might suggest the GP was considering a diagnosis of more serious illness but had not coded this. However, in those who received no relevant prescription in the first six months, there remained an increased risk of future diagnosis of COPD, asthma or IHD, and approximately half of all future diagnoses were made in such patients. This implies the possibility of a missed opportunity for more investigation of symptom causation and earlier initiation of treatment for COPD, asthma, and IHD.

Breathlessness symptom was associated with all-cause, COPD and IHD-related mortality in the short and long-term. All-cause mortality in the first six months after initial breathlessness symptom was particularly high and it is possible the symptom may be related to an underlying terminal disease. However, there were increased long-term risks for disease-specific mortality, which remained after adjustment for later diagnosis. This indicates that, in addition to being a marker of underlying diseases such as COPD, asthma and IHD, breathlessness remains an independent marker of long-term mortality risk.

*Comparison with previous literature*

In this study 12% of those with a newly recorded symptom of breathlessness went on to develop COPD and two community cohort studies found broadly similar results (7-15%) over 5 to 7 years with COPD diagnosis confirmed by spirometry.20,21 In a Swiss cohort 29% of those with shortness of breath followed over 11 years developed incident mild to moderate COPD.22 Another study using CPRD found that more than 20% of patients who were diagnosed with COPD had previously consulted with respiratory symptoms, and more than 40% diagnosed with COPD had already been prescribed a medication, in the 2 years before diagnosis.23 In our earlier case-control study, using a regional primary care database, 20% of those receiving a diagnosis of IHD had a prior code for breathlessness agreeing with our finding in CPRD of 17%.24

Breathlessness was found to be a significant predictor of long-term IHD mortality in a large Norwegian study,25 and a short-term predictor of cardiac arrest and myocardial infarction.26,27 In older people (≥70 years) in the UK community, breathlessness was found to be associated with all-cause mortality after adjustment for baseline COPD and left ventricular dysfunction.28

*Strengths and limitations*

This study used a large primary care database (CPRD) with previous validation of the identification of recording of COPD.29 The diagnosis recorded is that believed by the GP to be accurate at that time. Breathlessness symptom recorded in the consultation free text was not included. However, it is likely that those with a coded symptom are those with more troublesome symptoms of breathlessness, and those without an obvious reason for the symptom. No attempt was made to grade the degree of breathlessness.

*Implications for clinical practice*

Our findings lay out clearly and in quantitative terms the potential for more and detailed investigation, assessment and monitoring-over-time in order to earlier confirm or rule out COPD, asthma and IHD, and broad assessment of mortality risk in this common group of patients with breathlessness but no diagnosis. A cluster-randomised controlled trial suggested that targeted case-finding (inviting patients reporting relevant respiratory symptoms for post-bronchodilator spirometry) would improve the identification of people with COPD in primary care.30 In order to directly inform clinical decision-making and primary care policy, the next step is to establish whether realising this potential for early intervention in everyday practice diseases is feasible, benefits patients, and is cost-effective.

**Funding:** This paper presents independent research funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR) (grant reference number: SPCR-2014-10043). CCG is funded by the NIHR Collaborations for Leadership in Applied Health Research. The views expressed are those of the authors and not necessarily those of the NIHR, the NHS or the Department of Health.

**Ethical approval:** This study was approved by the CPRD Independent Scientific and Advisory Committee (ISAC, Protocol No. 15\_140).

**Competing interests:** None.

**Patient and public involvement:** Members from the North Staffordshire Breathe-Easy Group were invited to join a Patient Advisory Group. At the first meeting, the research questions and study design were discussed, at second and third meetings, views on the meaning of the results were sought, particularly how the results related to their own experiences, and to obtain advice on dissemination of the findings.

**Acknowledgments:** This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data is provided by patients and collected by the NHS as part of their care and support. Use of HES Data/ONS Data: Copyright © (2018), re-used with the permission of The Health & Social Care Information Centre. All rights reserved. The interpretation and conclusions contained in this study are those of the authors alone.

The authors would like to thank the members from the North Staffordshire Breathe-Easy Group who participated in this study as a Patient Advisory Group.

**References**

1. Herman L, Atri A, Salloway S. Alzheimer’s disease in primary care: the significance of early detection, diagnosis, and intervention. *Am J Med* 2017; 130(6): 756.

2. Feldman DE, Bernatsky S, Houde M, Beauchamp ME, Abrahamowicz M. Early consultation with a rheumatologist for RA: does it reduce subsequent use of orthopaedic surgery? *Rheumatology (Oxford)* 2013; 52(3): 452-459.

3. Heuvelings CC, de Vries SG, Greve PF, *et al*. Effectiveness of interventions for diagnosis and treatment of tuberculosis in hard-to-reach populations in countries of low and medium tuberculosis incidence: a systematic review. *Lancet Infect Dis* 2017; 17(5): e144-158.

4. Muchtar E, Gertz MA, Kumar SK, *et al*. Improved outcomes for newly diagnosed AL amyloidosis between 2000 and 2014: cracking the glass ceiling of early death. *Blood* 2017; 129(15): 2111-2119.

5. Jones R. Can early diagnosis and effective management combat the irresistible rise of COPD? *Br J Gen Pract* 2006; 56(530): 652-654.

6. [Lindberg A](https://www.ncbi.nlm.nih.gov/pubmed/?term=Lindberg%20A%5BAuthor%5D&cauthor=true&cauthor_uid=15888826), [Jonsson AC](https://www.ncbi.nlm.nih.gov/pubmed/?term=Jonsson%20AC%5BAuthor%5D&cauthor=true&cauthor_uid=15888826), [Rönmark E](https://www.ncbi.nlm.nih.gov/pubmed/?term=R%C3%B6nmark%20E%5BAuthor%5D&cauthor=true&cauthor_uid=15888826), [Lundgren R](https://www.ncbi.nlm.nih.gov/pubmed/?term=Lundgren%20R%5BAuthor%5D&cauthor=true&cauthor_uid=15888826), [Larsson LG](https://www.ncbi.nlm.nih.gov/pubmed/?term=Larsson%20LG%5BAuthor%5D&cauthor=true&cauthor_uid=15888826), [Lundbäck B](https://www.ncbi.nlm.nih.gov/pubmed/?term=Lundb%C3%A4ck%20B%5BAuthor%5D&cauthor=true&cauthor_uid=15888826). Ten-year cumulative incidence of COPD and risk factors for incident disease in a symptomatic cohort. *Chest* 2005; 127(5): 1544-1552.

7. Løkke A, Lange P, Scharling H, Fabricius P, Vestbo J. Developing COPD: a 25 year follow up study of the general population. *Thorax* 2006; 61(11): 935-939.

8. [Pirie K](https://www.ncbi.nlm.nih.gov/pubmed/?term=Pirie%20K%5BAuthor%5D&cauthor=true&cauthor_uid=23107252), [Peto R](https://www.ncbi.nlm.nih.gov/pubmed/?term=Peto%20R%5BAuthor%5D&cauthor=true&cauthor_uid=23107252), [Reeves GK](https://www.ncbi.nlm.nih.gov/pubmed/?term=Reeves%20GK%5BAuthor%5D&cauthor=true&cauthor_uid=23107252), [Green J](https://www.ncbi.nlm.nih.gov/pubmed/?term=Green%20J%5BAuthor%5D&cauthor=true&cauthor_uid=23107252), [Beral V](https://www.ncbi.nlm.nih.gov/pubmed/?term=Beral%20V%5BAuthor%5D&cauthor=true&cauthor_uid=23107252); [Million Women Study Collaborators](https://www.ncbi.nlm.nih.gov/pubmed/?term=Million%20Women%20Study%20Collaborators%5BCorporate%20Author%5D). The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. *Lancet* 2013; 381(9861): 133-141.

9. Loymans RJ, Gemperli A, Cohen J, *et al*. Comparative effectiveness of long term drug treatment strategies to prevent asthma exacerbations: network meta-analysis. *BMJ* 2014; 348: g3009.

10. Stuart-Buttle CD, Read JD, Sanderson HF, Sutton YM. A language of health in action: Read Codes, classifications and groupings. *Proc AMIA Annu Fall Symp* 1996: 75-79.

11. Quality and Outcomes Framework. Available at <http://content.digital.nhs.uk/qof> Accessed 17 April 2019.

12. [Herrett E](https://www.ncbi.nlm.nih.gov/pubmed/?term=Herrett%20E%5BAuthor%5D&cauthor=true&cauthor_uid=26050254), [Gallagher AM](https://www.ncbi.nlm.nih.gov/pubmed/?term=Gallagher%20AM%5BAuthor%5D&cauthor=true&cauthor_uid=26050254), [Bhaskaran K](https://www.ncbi.nlm.nih.gov/pubmed/?term=Bhaskaran%20K%5BAuthor%5D&cauthor=true&cauthor_uid=26050254), *et al*. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015; 44(3): 827-836.

13. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010; 69(1): 4-14.

14. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systemic review. *Br J Gen Pract* 2010; 60(572): e128-136.

15. Gallagher AM, Puri S, Van Staa T. Linkage of the General Practice Research Database (GPRD) with other data sources. *Pharmacoepidemiol Drug Saf* 2011; 20: S230-231.

16. English Indices of Deprivation 2010. Available at <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2010> Accessed 17 April 2019.

17. Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. *Fam Pract* 2013; 30(2): 172-178.

18. Gibson PG, Simpson JL. The overlap syndrome of asthma and COPD: what are its features and how important is it? *Thorax* 2009; 64(8): 728-735.

19. [Bateman ED](https://www.ncbi.nlm.nih.gov/pubmed/?term=Bateman%20ED%5BAuthor%5D&cauthor=true&cauthor_uid=26255108), [Reddel HK](https://www.ncbi.nlm.nih.gov/pubmed/?term=Reddel%20HK%5BAuthor%5D&cauthor=true&cauthor_uid=26255108), [van Zyl-Smit RN](https://www.ncbi.nlm.nih.gov/pubmed/?term=van%20Zyl-Smit%20RN%5BAuthor%5D&cauthor=true&cauthor_uid=26255108), [Agusti A](https://www.ncbi.nlm.nih.gov/pubmed/?term=Agusti%20A%5BAuthor%5D&cauthor=true&cauthor_uid=26255108). The asthma-COPD overlap syndrome: towards a revised taxonomy of chronic airways diseases? *Lancet Respir Med* 2015; 3(9): 719-728.

20. Vestbo J, Lange P. Can GOLD Stage 0 provide information of prognostic value in chronic obstructive pulmonary disease? *Am J Respir Crit Care Med* 2002; 166(3): 329-332.

21. Lindberg A, Eriksson B, Larsson LG, Rönmark E, Sandström T, Lundbäck B. Seven-year cumulative incidence of COPD in an age-stratified general population sample. *Chest* 2006; 129(4): 879-885.

22. Probst-Hensch NM, Curjuric I, Pierre-Olivier B, *et al*. Longitudinal change of prebronchodilator spirometric obstruction and health outcomes: results from the SAPALDIA cohort. *Thorax* 2010; 65(2): 150-156.

23. Jones RC, Price D, Ryan D, *et al*. Opportunities to diagnose chronic obstructive pulmonary disease in routine care in the UK: a retrospective study of a clinical cohort. *Lancet Respir Med* 2014; 2(4): 267-276.

24. Hayward RA, Chen Y, Croft P, Jordan KP. Presentation of respiratory symptoms prior to diagnosis in general practice: a case–control study examining free text and morbidity codes. *BMJ Open* 2015; 5(6): e007355.

25. Frostad A, Soyseth V, Haldorsen T, Andersen A, Gulsvik A. Respiratory symptoms and long-term cardiovascular mortality. *Respir Med* 2007; 101(11): 2289-2296.

26. Höglund H, Jansson JH, Forslund AS, Lundblad D. Prodromal symptoms and health care consumption prior to out-of-hospital cardiac arrest in patients without previously known ischaemic heart disease. *Resuscitation* 2014; 85(7): 864-868.

27. Sequist TD, Marshall R, Lampert S, Buechler EJ, Lee TH. Missed opportunities in the primary care management of early acute ischemic heart disease. *Arch Intern Med* 2006; 166(20): 2237-2243.

28. Ahmed T, Steward JA, O’Mahony MS. Dyspnoea and mortality in older people in the community: a 10-year-up. *Age Ageing* 2012; 41(4): 545-549.

29. Quint JK, Müllerova H, DiSantostefano Rl, *et al*. Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD). *BMJ Open* 2014; 4(7): e005540.

30. Jordan RE, Adab P, Sitch A, *et al*. Targeted case finding for chronic obstructive pulmonary disease versus routine practice in primary care (TargetCOPD): a cluster-randomised controlled trial. *Lancet Respir Med* 2016; 4(9): 720-730.

Heading

Figure 1 Kaplan-Meier curves for first diagnosis of COPD, asthma and IHD for the ‘exposed’ and ‘unexposed’ cohorts

Legends

Red, the ‘exposed’ cohort; Blue, the ‘unexposed’ cohort. At least six months’ follow-up was available for 93% of patients, and 12% of patients had the full 12 years of follow-up (maximum follow-up time, 18 years). Kaplan-Meier curves truncated at 12 years.



**Table 1 Participant demographic and clinical characteristics at baseline**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Patients with a first coded record of breathlessness symptom | Matched participants |  |
| Participants, *n* | 75,698 | 75,698 | matched |
| Year of index date, median (IQR) | 2005 (2001, 2007) | 2005 (2001, 2007) | matched |
| Age, *n* (%) |  |  | matched |
|  | 18 - 44 | 16,668 (22.0) | 16,668 (22.0) |  |
|  | 45 - 59 | 20,356 (26.9) | 20,356 (26.9) |  |
|  | 60 - 74 | 23,813 (31.5) | 23,813 (31.5) |  |
|  | 75 and over | 14,861 (19.6) | 14,861 (19.6) |  |
| Male, *n* (%) | 32,195 (42.5) | 32,195 (42.5) | matched |
| Geographical region, *n* (%) |  |  | matched |
|  | London  | 6,154 (8.1) | 6,154 (8.1) |  |
|  | South | 19,592 (25.9) | 19,592 (25.9) |  |
|  | Midlands and East | 24,219 (32.0) | 24,219 (32.0) |  |
|  | North | 25,733 (34.0) | 25,733 (34.0) |  |
| Smoking, *n* (%) |  |  | *p* < 0.001 |
|  | Non smoker | 40,436 (53.4) | 42,632 (56.3) |  |
|  | Ever smoker | 24,740 (32.7) | 19,143 (25.3) |  |
|  | Unknown | 10,522 (13.9) | 13,923 (18.4) |  |
| Alcohol consumption, *n* (%) |  |  | *p* < 0.001 |
|  | Non drinker | 8,623 (11.4) | 7,425 (9.8) |  |
|  | Ever drinker | 54,991 (72.7) | 52,197 (69.0) |  |
|  | Unknown | 12,084 (16.0) | 16,076 (21.2) |  |
| Body mass index, *n* (%)  |  |  | *p* < 0.001 |
|  | < 25 kg/m2 | 23,901 (31.6) | 27,054 (35.7) |  |
|  | ≥ 25 kg/m2 | 39,854 (52.7) | 32,194 (42.5) |  |
|  | Unknown | 11,943 (15.8) | 16,450 (21.7) |  |
| Deprivation level, *n* (%) |  |  | *p* < 0.001 |
|  | 1 (least) | 17,008 (22.5) | 18,231 (24.1) |  |
|  | 2 | 18,123 (23.9) | 19,043 (25.2) |  |
|  | 3 | 15,810 (20.9) | 15,644 (20.7) |  |
|  | 4 | 13,508 (17.8) | 12,772 (16.9) |  |
|  | 5 (most) | 11,163 (14.8) | 9,929 (13.1) |  |
|  | Unknown | 86 (0.1) | 79 (0.1) |  |
| Number of co-morbidities, median (IQR) | 6 (3, 11) | 3 (1, 7) | *p* < 0.001 |

IQR, interquartile range; *p*-value obtained from Chi-squared or from Mann-Whitney U test as appropriate, and where applicable analysis excluded the unknown category (missing data).

**Table 2 Incidence of future diagnosis of COPD, asthma or IHD in patients with a first coded record of breathlessness symptom and in the matched participants**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Patients with a first coded record of breathlessness symptom (*n*=75,698) | Matched participants (*n*=75,698) | Patients vs. matched participants |
| Time period | Number with diagnosis (%\*) | Time at risk (person-years) | Incidence† | Number with diagnosis (%\*) | Time at risk (person-years) | Incidence† | Adjusted‡ hazard ratio (95% CI) |
| COPD |  |  |  |  |  |  |  |
| Diagnostic window (0-6m) | 2,970 (3.9) | 35349.8  | 840.2 | 72 (0.1) | 37,188.2  | 19.4 | - |
| Long term follow-up (>6m) | 6,141 (8.1) | 441,810.8 | 139.0 | 1,883 (2.5) | 503,290.6 | 37.4 | - |
|  | 6–12m | 685 (0.9) | 33,467.6  | 204.7 | 80 (0.1) | 36,235.7  | 22.1 | 8.6 (6.8, 11.0) |
|  | 12–36m | 1,835 (2.4) | 120,124.0  | 152.8 | 364 (0.5) | 133,356.3  | 27.3 | 5.3 (4.7, 6.0) |
|  | >36m | 3,621 (4.8) | 288,219.2  | 125.6 | 1,439 (1.9) | 333,698.6  | 43.1 | 2.8 (2.6, 3.0) |
| Asthma |  |  |  |  |  |  |  |
| Diagnostic window (0-6m) | 4,430 (5.9) | 34,759.7 | 1274.5 | 80 (0.1) | 37,184.6  | 21.5 | - |
| Long term follow-up (>6m) | 6,925 (9.1) | 411,223.4 | 168.4 | 1,347 (1.8) | 502,628.2 | 26.8 | - |
|  | 6–12m | 1,032 (1.4) | 32,539.8  | 317.2 | 89 (0.1) | 36,226.2  | 24.6 | 11.7 (9.4, 14.6) |
|  | 12–36m | 2,335 (3.1) | 115,176.6  | 202.7 | 347 (0.5) | 133,251.3  | 26.0 | 6.8 (6.0, 7.6) |
|  | >36m | 3,558 (4.7) | 263,507.0  | 135.0 | 911 (1.2) | 333,150.7  | 27.3 | 4.3 (3.9, 4.6) |
| IHD |  |  |  |  |  |  |  |
| Diagnostic window (0-6m) | 4,430 (5.9) | 34,964.4  | 1267.0 | 338 (0.4) | 37,131.6  | 91.0 | - |
| Long term follow-up (>6m) | 8,680 (11.5) | 427,498.9 | 203.0 | 5,240 (6.9) | 489,700.3 | 107.0 | - |
|  | 6–12m | 1,115 (1.5) | 32,864.5  | 339.3 | 386 (0.5) | 36,068.3  | 107.0 | 3.0 (2.7, 3.4) |
|  | 12–36m | 2,562 (3.4) | 117,374.1  | 218.3 | 1,390 (1.8) | 131,714.2  | 105.5 | 1.9 (1.8, 2.1) |
|  | >36m | 5,003 (6.6) | 277,260.3  | 180.4 | 3,464 (4.6) | 321,917.8  | 107.6 | 1.6 (1.5, 1.7) |

\*Percentage in the total number of patients with a first coded record of breathlessness symptom; or in the total number of the matched participants. †Per 10,000 person-years. ‡Adjusted hazard ratio, adjusted for index year, age, gender, region, ever smoking, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain.

**Table 3 Prescription in first six months after index date in patients with a first coded record of breathlessness symptom**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | COPD | Asthma | IHD |
|  |  |  | No. of patients | No. (%) who received prescription\* in first 6 months | No. of patients | No. (%) who received prescription\* in first 6 months | No. of patients | No. (%) who received prescription\* in first 6 months |
| Medication prescribed† |  |  |  |  |  |  |
| In whole cohort | 75,698 | 26,055 (34.4) | 75,698 | 26,313 (34.8) | 75,698 | 25,828 (34.1) |
|  | In those received a diagnosis‡ | 9,111 | 4,810 (52.8) | 11,355 | 8,080 (71.2) | 13,110 | 7360 (56.1) |
|  |  | Diagnosis during 0-6m | 2,970 | 1,969 (66.3) | 4,430 | 3,918 (88.4) | 4,430 | 2,398 (54.1) |
|  |  | Diagnosis during 6-12m | 685 | 390 (56.9) | 1,032 | 784 (76.0) | 1,115 | 744 (66.7) |
|  |  | Diagnosis during 12-36m | 1,835 | 878 (47.8) | 2,335 | 1,511 (64.7) | 2,562 | 1,591 (62.1) |
|  |  | Diagnosis after 36m | 3,621 | 1,573 (43.4) | 3,558 | 1,867 (52.5) | 5,003 | 2,627 (52.5) |
|  | In those did not receive a diagnosis‡ | 66,587 | 21,245 (31.9) | 64,343 | 18,233 (28.3) | 62,588 | 18,468 (29.5) |

\*Before the date of diagnosis if diagnosis occurred within first 6 months (pre-diagnosis prescription). †COPD/asthma-related medications regarding COPD and asthma; and IHD-related medications regarding IHD. ‡Diagnosis during the follow-up period.

**Table 4 Risk of future diagnosis of COPD, asthma and IHD in patients with a first coded record of breathlessness symptom stratified by prescription management**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Time period | Patient group\* | COPDAdjusted† hazard ratio (95% CI) | AsthmaAdjusted† hazard ratio (95% CI) | IHDAdjusted† hazard ratio (95% CI) |
|  | A | 1.0 (referent) | 1.0 (referent) | 1.0 (referent) |
| 6–12m | B | 5.6 (4.3, 7.2) | 4.4 (3.5, 5.7) | 2.0 (1.8, 2.4) |
|  | C | 16.9 (13.1, 21.7) | 29.0 (23.1, 36.3) | 4.1 (3.6, 4.7) |
|  |  |  |  |  |
|  | A | 1.0 (referent) | 1.0 (referent) | 1.0 (referent) |
| 12–36m | B | 4.0 (3.6, 4.6) | 3.7 (3.2, 4.2) | 1.5 (1.4, 1.6) |
|  | C | 8.8 (7.7, 10.0) | 14.6 (12.9, 16.5) | 2.4 (2.2, 2.6) |
|  |  |  |  |  |
|  | A | 1.0 (referent) | 1.0 (referent) | 1.0 (referent) |
| >36m | B | 2.3 (2.1, 2.5) | 3.0 (2.7, 3.2) | 1.4 (1.3, 1.4) |
|  | C | 4.1 (3.8, 4.4) | 7.7 (7.1, 8.4) | 2.0 (1.9, 2.1) |

\*Patient group: A – matched (‘unexposed’) participants; B - patients with a first coded record of breathlessness symptom and with no relevant prescribing in first six months after index date and before diagnosis (if any); C - patients with a first coded record of breathlessness symptom and with relevant prescribing in first six months after index date and before diagnosis (if any); Size of patient group: A (*n*=75,698); B (*n*=49,643 for COPD; *n*=49,385 for asthma; *n*=49,870 for IHD); C (*n*=26,055 for COPD; *n*=26,313 for asthma; *n*=25,828 for IHD). †Adjusted hazard ratio, adjusted for index year, age, gender, region, ever smoking, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain.

**Table 5 Risk of mortality in patients with a first coded record of breathlessness symptom, compared to the matched participants**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Time period | Patients with a first coded record of breathlessness symptom | Matched participants | Patients vs. matched participants |
|  |  | Number of events (%) | Time at risk (person-years) | Incidence\* | Number of events (%) | Time at risk (person-years) | Incidence\* | Adjusted† hazard ratio‡ (95% CI) |
| All-cause | 0–6m | 3,119 (4.1) | 35,962.2 | 867.3 | 423 (0.6) | 36,801.9 | 114.9 | 6.0 (5.4, 6.7) |
| >6m | 11,251 (14.9) | 462,060.6 | 243.5 | 7,726 (10.2) | 489,616.9 | 157.8 | 1.2 (1.1, 1.3) |
|  |  |  |  |  |  |  |  |  |
| COPD-related | 0–6m | 78 (0.1) | 35,962.2 | 21.7 | 5 (0.0) | 36,801.9 | 1.4 | 11.5 (4.5, 28.9) |
| >6m | 1,030 (1.4) | 462,060.6 | 22.3 | 198 (0.3) | 489,616.9 | 4.0 | 1.8 (1.5, 2.1) |
|  |  |  |  |  |  |  |  |  |
| IHD-related | 0–6m | 452 (0.6) | 35,962.2 | 125.7 | 70 (0.1) | 36,801.9 | 19.0 | 4.7 (3.6, 6.2) |
| >6m | 2,999 (4.0) | 462,060.6 | 64.9 | 1,625 (2.1) | 489,616.9 | 33.2 | 1.2 (1.1, 1.3) |

\*Per 10,000 person-years. †Adjusted for index year, age, gender, region, ever smoking, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain and later diagnosis; Later diagnosis as a time-varying covariate; For all-cause mortality, later diagnosis is the earliest diagnosis of asthma, COPD or IHD. ‡Sub-distribution hazard ratio was used in completing risk models for disease-specific mortality (other death was treated as competing events); and hazard ratio in Cox models for all-cause mortality.

**Supplementary Table 1 Other covariates (co-morbid conditions) measured at baseline**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Patients with a first coded record of breathlessness symptom | Matched participants |  |
| Participants, *n* | 75,698 | 75,698 | matched |
| Respiratory symptom, *n* (%) |  |  |  |
|  | Cough | 18,099 (23.9) | 6,149 (8.1) | *p* < 0.001 |
|  | Sputum | 1,206 (1.6) | 305 (0.4) | *p* < 0.001 |
| Inflammatory condition, *n* (%) |  |  |  |
|  | Rheumatoid arthritis | 479 (0.6) | 261 (0.3) | *p* < 0.001 |
|  | Gout | 979 (1.3) | 707 (0.9) | *p* < 0.001 |
|  | Polymyalgia rheumatic | 544 (0.7) | 298 (0.4) | *p* < 0.001 |
|  | Inflammatory bowel disease | 294 (0.4) | 183 (0.2) | *p* < 0.001 |
|  | Systemic lupus erythematosus | 24 (0.03) | 11 (0.01) | *p* = 0.028 |
|  | Spondyloarthritis\* | 153 (0.2) | 79 (0.1) | *p* < 0.001 |
| Mental condition, *n* (%) |  |  |  |
|  | Depression | 4,176 (5.5) | 2,274 (3.0) | *p* < 0.001 |
|  | Anxiety  | 4,513 (6.0) | 2,479 (3.3) | *p* < 0.001 |
| Diabetes, *n* (%) | 4,780 (6.3) | 3,668 (4.9) | *p* < 0.001 |
| Non-specific chest pain, *n* (%) | 6,335 (8.4) | 2,343 (3.1) | *p* < 0.001 |
| Musculoskeletal pain, *n* (%) | 23,020 (30.4) | 15,795 (20.9) | *p* < 0.001 |

\*Including ankylosing spondylitis and psoriatic arthritis. *p* value obtained from Chi-square.

**Supplementary Table 2 Incidence of future diagnosis of COPD, asthma or IHD\* in patients with a first coded record of breathlessness symptom and in the matched participants**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Patients with a first coded record of breathlessness symptom*n*=75,698 | Matched participants*n*=75,698 | Patients vs. matched participants |
| Time period | Number with diagnosis (%†) | Time at risk (person-years) | Incidence‡  | Number with diagnosis (%†) | Time at risk (person-years) | Incidence‡ | Adjusted§ hazard ratio (95% CI) |
| Diagnostic window (0-6m) | 11,150 (14.7) | 32,558.2 | 3,424.6 | 474 (0.6) | 37,100.0 | 127.8 | - |
| Long term follow-up (>6m) | 15,669 (20.7) | 345,585.7 | 453.4 | 7,440 (9.8) | 479,856.2 | 155.0 | - |
|  | 6–12m | 2,411 (3.2) | 29,367.1 | 821.0 | 532 (0.7) | 35,971.7 | 147.9 | 5.1 (4.7, 5.7) |
|  | 12–36m | 5,212 (6.9) | 100,528.8 | 518.5 | 1,959 (2.6) | 130,733.8 | 149.8 | 3.2 (3.0, 3.3) |
|  | >36m | 8,046 (10.6) | 215,689.8 | 373.0 | 4,949 (6.5) | 313,150.7 | 158.0 | 2.2 (2.1, 2.3) |

\*First diagnosed condition if multiple conditions occurred. †Percentage in the total number of patients with a first coded record of breathlessness symptom; or in the total number of the matched participants. ‡Per 10,000 person-years. §Adjusted hazard ratio, adjusted for index year, age, gender, region, ever smoking, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain.

**Supplementary Table 3 Incidence of future diagnosis of COPD, asthma or IHD\* in patients with a first coded record of breathlessness symptom and in the matched participants, stratified by smoking status**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Patients with a first coded record of breathlessness symptom | Participants from the matched cohort | Patients vs. participants |
| Time period | Number with diagnosis (%†) | Time at risk (person-years) | Incidence‡  | Number with diagnosis (%†) | Time at risk (person-years) | Incidence‡ | Adjusted§ hazard ratio (95% CI) |
| Ever-smoking subset | *n*=24,740 | *n*=19,143 |  |
| Diagnostic window (0–6m) | 4,493 (18.2) | 10,284.6 | 4368.7 | 154 (0.8) | 9,376.7 | 164.2 | - |
| Long term follow-up (>6m) | 5,821 (23.5) | 99,510.7 | 585.0 | 2,375 | 116,093.4 | 204.6 | - |
|  | 6–12m | 955 (3.9) | 9,090.2 | 1050.6 | 167 (0.9) | 9,069.0 | 184.1 | 5.58 (4.71, 6.61) |
|  | 12–36m | 2,003 (8.1) | 30,364.2 | 659.7 | 618 (3.2) | 32,711.2 | 188.9 | 3.41 (3.10, 3.74) |
|  | >36m | 2,863 (11.6) | 60,056.3 | 476.7 | 1,590 (8.3) | 74,313.2 | 214.0 | 2.18 (2.0, 2.33) |
|  |  |  |  |  |  |  |  |  |
| Never-smoking subset | *n*=40,436 | *n*=42,632 |  |
| Diagnostic window (0–6m) | 4,995 (12.4) | 17,820.8 | 2802.9 | 227 (0.5) | 20,918.6 | 108.5 | - |
| Long term follow-up (>6m) | 7,481 (18.5) | 197,099.7 | 379.6 | 3,695 (8.7) | 272,711.4 | 135.5 | - |
|  | 6–12m | 1,118 (2.8) | 16,285.1 | 686.5 | 270 (0.6) | 20,336.7 | 132.8 | 4.77 (4.16, 5.47) |
|  | 12–36m | 2,429 (6.0) | 56,565.6 | 429.4 | 1,012 (2.4) | 74,332.1 | 136.1 | 2.84 (2.63, 3.07) |
|  | >36m | 3,934 (9.7) | 124,249.0 | 316.6 | 2,413 (5.7) | 178,042.6 | 135.5 | 2.18 (2.07, 2.30) |

\*First diagnosed condition if multiple conditions occurred. †Percentage in the total number of ever-smoking patients with a first coded record of breathlessness symptom; or in the total number of ever-smoking participants from the matched cohort. ‡Per 10,000 person-years. §Adjusted hazard ratio, adjusted for index year, age, gender, region, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain. Total number of patients different to 75,698 in each cohort was due to missing data in smoking status.

**Supplementary Table 4 Risk of mortality in patients with a first coded record of breathlessness symptom compared to the matched participants stratified by age group**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Time period | Patients with a first coded record of breathlessness symptom | Matched participants | Patients vs. matched participants |
|  |  | Number of events (%) | Time at risk (person-years) | Incidence\* | Number of events (%) | Time at risk (person-years) | Incidence\* | Adjusted† hazard ratio‡ (95% CI) |
| Aged 18-59 years |  |  | *n*=37,024 |  |  | *n*=37,024 |  |  |
| All-cause | 0–6m | 512 (1.4) | 17,896.9 | 286.1 | 30 (0.08) | 18,001.4 | 16.7 | 11.7 (8.1, 17.0) |
| >6m | 1,467 (4.0) | 251,019.6 | 58.4 | 685 (1.9) | 253,278.6 | 27.0 | 1.4 (1.3, 1.6) |
| COPD-related | 0–6m | 4 (0.01) | 17,896.9 | 2.2 | 0 (0.0) | 18,001.4 | 0.0 | - |
| >6m | 117 (0.3) | 251,019.6 | 4.7 | 11 (0.03) | 253,278.6 | 0.4 | 2.8 (1.4, 5.6) |
| IHD-related | 0–6m | 37 (0.1) | 17,896.9 | 20.7 | 2 (0.01) | 18,001.4 | 1.1 | 14.7 (3.5, 61.7) |
| >6m | 239 (0.6) | 251,019.6 | 9.5 | 105 (0.3) | 253,278.6 | 4.1 | 1.6 (1.3, 2.1) |
|  |  |  |  |  |  |  |  |  |
| Aged 60 years and above |  |  | *n*=38,674 |  |  | *n*=38,674 |  |  |
| All-cause | 0–6m | 2,607 (6.7) | 18,065.4 | 1,443.1 | 393 (1.0) | 18,800.5 | 209.0 | 5.6 (5.0, 6.2) |
| >6m | 9,784 (25.3) | 211,040.9 | 463.6 | 7,041 (18.2) | 236,338.3 | 297.9 | 1.2 (1.1, 1.3) |
| COPD-related | 0–6m | 74 (0.2) | 18,065.4 | 41.0 | 5 (0.01) | 18,800.5 | 2.7 | 9.8 (3.9, 24.5) |
| >6m | 913 (2.4) | 211,040.9 | 43.3 | 187 (0.5) | 236,338.3 | 7.9 | 1.6 (1.3, 1.9) |
| IHD-related | 0–6m | 415 (1.1) | 18,065.4 | 229.7 | 68 (0.2) | 18,800.5 | 36.2 | 4.7 (3.6, 6.1) |
| >6m | 2,760 (7.1) | 211,040.9 | 130.8 | 1,520 (3.9) | 236,338.3 | 64.3 | 1.2 (1.1, 1.3) |

\*Per 10,000 person-years. †Adjusted for index year, age, gender, region, ever smoking, ever alcohol drinking, BMI, deprivation level, comorbidity, cough and sputum, inflammatory conditions, depression and anxiety, diabetes, non-specific chest pain, musculoskeletal pain and later diagnosis; Later diagnosis as a time-varying covariate; For all-cause mortality, later diagnosis is the earliest diagnosis of asthma, COPD or IHD. ‡Sub-distribution hazard ratio was used in completing risk models for disease-specific mortality (other death was treated as competing events); and hazard ratio in Cox models for all-cause mortality.